This paper must be cited as:


The final publication is available at

[http://dx.doi.org/10.1016/j.jip.2016.11.014]

Copyright [Elsevier]
Title: Novel RNA viruses producing simultaneous covert infections in Ceratitis capitata. Correlations between viral titers and host fitness, and implications for SIT programs

Article Type: Full length article

Keywords: iflavirus; noravirus; covert infection; persistent infection; medfly

Corresponding Author: Dr. Salvador Herrero Sendra, PhD

Corresponding Author's Institution: University of Valencia

First Author: Angel Llopis-Giménez

Order of Authors: Angel Llopis-Giménez; Rosa M González-Martínez; Anabel Millán-Leiva; Marta Catalá; Elena Llacer; Alberto Urbaneja; Salvador Herrero Sendra, PhD

Abstract: The Mediterranean fruit fly (medfly), Ceratitis capitata is a highly polyphagous pest, which infests multiple species of fruits and vegetables worldwide. In addition to the traditional control with chemical insecticides, sterile insect technique (SIT) has been implemented in integrated programs worldwide, and has become an essential measure for the control of this pest. A key issue for SIT is to release sterile males that are sufficiently competitive with males from the wild population. Using sequence information available in public databases, three novel picornaviruses infecting medflies were discovered and named as C. capitata iflavirus 1 and 2 (CcaIV1 and CcaIV2), and C. capitata noravirus (CcaNV). Additional analyses have revealed the presence of CcaIV2 and CcaNV covertly infecting most of the medfly strains used in the different SIT programs around the world, as well as in field captures in the east of Spain. High viral titers of CcaNV were associated with a reduction in the lifespan of males released to the field for the control of this pest, suggesting the possibility that CcaNV may impair the fitness of sterile flies produced by SIT programs.
Dear Editor,

Please find attached a revised version of our manuscript entitled: “Novel RNA viruses producing simultaneous covert infections in Ceratitis capitata. Correlations between viral titers and host fitness, and implications for SIT programs” including the changes you have suggested.

We understand and share your concerns about the fitness data and we are glad you have found a way to maintaining the fitness data on the final version of the manuscript. We also appreciate your suggestions and effort to address it in a more conservative way.

We have highlighted in red the main changes in the new version of the manuscript.

Sincerely

Salva Herrero
No needed
Medfly pupae for SIT
• Three novel RNA viruses have been found covertly infecting medflies

• Sterile insect technique (SIT) has been implemented in integrated programs for medfly control.

• Two of these viruses infects field as well as insects used for the mass-production of males for SIT programs

• Higher viral titers of CcaNV have been found in SIT males with shorter lifespan
Novel RNA viruses producing simultaneous covert infections in *Ceratitis capitata*. Correlations between viral titers and host fitness, and implications for SIT programs

**Authors:** Angel Llopis-Giménez\(^1,2\)#, Rosa María González\(^1,2\)#, Anabel Millán-Leiva\(^1,2\), Marta Catalá\(^3\), Elena Llacer\(^3\), Alberto Urbaneja\(^3\), and Salvador Herrero\(^1,2\)*

**Affiliations:**

\(^1\)Department of Genetics, Universitat de València, Dr Moliner 50, 46100 Burjassot, Spain

\(^2\)Estructura de Recerca Interdisciplinar en Biotecnologia i Biomedicina (ERI-BIOTECMED), Universitat de València. Dr Moliner 50, 46100 Burjassot, Spain.

\(^3\)Instituto Valenciano de Investigaciones Agrarias. Unidad Asociada de Entomología IVIA-UJI. Centro de Protección Vegetal y Biotecnología, 46113-Moncada, España.

#Contributed equally to this work

*Correspondence to: Salvador Herrero
Department of Genetics
Universitat de València
Dr Moliner 50, 46100 Burjassot, Spain
sherrero@uv.es
Phone: + 34 96 354 30 06
Fax: + 34 96 354 30 29

**One Sentence Summary:** Viral covert infections in medfly

**Keywords:** iflavirus, noravirus, covert infection, persistent infection, medfly
Abstract

The Mediterranean fruit fly (medfly), *Ceratitis capitata* is a highly polyphagous pest, which infests multiple species of fruits and vegetables worldwide. In addition to the traditional control with chemical insecticides, sterile insect technique (SIT) has been implemented in integrated programs worldwide, and has become an essential measure for the control of this pest. A key issue for SIT is to release sterile males that are sufficiently competitive with males from the wild population. Using sequence information available in public databases, three novel picornaviruses infecting medflies were discovered and named as *C. capitata* iflaviruses 1 and 2 (CcaIV1 and CcaIV2), and *C. capitata* noravirus (CcaNV). Additional analyses have revealed the presence of CcaIV2 and CcaNV covertly infecting most of the medfly strains used in the different SIT programs around the world, as well as in field captures in the east of Spain. High viral titers of CcaNV were associated with a reduction in the lifespan of males released to the field for the control of this pest, suggesting the possibility that CcaNV may impair the fitness of sterile flies produced by SIT programs.
1. Introduction

The Mediterranean fruit fly, or medfly, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae), is a highly polyphagous species whose larvae develop in a very wide range of unrelated fruits belonging to a large number of families. Practically all tree fruit crops might be a host for this threatening fruit fly pest and, indeed, hosts of economic importance include *Citrus* spp., apples (*Malus pumila*), avocados (*Persea americana*), figs (*Ficus carica*), kiwifruits (*Actinidia deliciosa*), mangoes (*Mangifera indica*), medlars (*Mespilus germanica*), pears (*Pyrus communis*), and *Prunus* spp. (especially peaches, *P. persica*) ((EPPO), 2016).

This wide range of important hosts makes the medfly one of the world’s most damaging fruit pests which causes serious economic losses worldwide (Ismay, 1992). In the Mediterranean citrus growing areas, *C. capitata* is considered a major citrus pest due to massive fruit losses and quarantine restrictions (Enkerlin and Mumford, 1997; Tena and Garcia-Marí, 2011).

Traditionally, medfly management has been mainly based on field monitoring and aerial and terrestrial treatments with organophosphate insecticides mixed with proteinaceous baits (Andres et al., 2009; Chueca et al., 2007). To reduce the environmental impacts of these treatments, during recent decades, sterile insect technique (SIT) programs have been implemented in North, Central, and South America, Europe, the Middle East, Asia, Africa, and Australia (Enkerlin, 2005), and have become an essential measure to control the medfly worldwide. Currently, the different fruit fly factories mass rear several billion sterile medfly males weekly that are released in the field to mate with wild females. Thus, a key issue for SIT is to release sterile males that are sufficiently competitive with males from the wild population (Andres et al., 2009; Juan-Blasco et al., 2013). Therefore, research during recent years has addressed sterile male mating competitiveness to implement cost-effective SIT programs (Ami et al., 2009; Andres et al., 2009; Juan-Blasco et al., 2013; Pereira et al., 2013).
Insect-infecting picornaviruses (Order *Picornavirales*) belong to two families, *Iflaviridae* and *Dicistroviridae*. In addition, a few unclassified picornaviruses described in *Drosophila* spp. (Diptera: Drosophilidae) and the lepidopteran *Spodoptera exigua* (Hübner) (Lepidoptera: Noctuidae) may represent a new family within this order. Those viruses form non-enveloped and icosahedral particles that contain a single copy of a single-stranded RNA genome. For the iflaviruses, the genome contains a single open reading frame (ORF) that is translated into a large polypeptide that is processed into functional viral proteins. In the case of dicistroviruses, the genome is bicistronic, encoding one polypeptide for the non-structural proteins and a second polypeptide for the structural proteins (King, 2011; Possee and King, 2001). In the case of *C. capitata*, the few reports of RNA viruses (one of them morphologically resembling picornavirus particles) date to the beginning of the 1980s (Plus et al., 1981).

The pathogenicity of insect picornaviruses can vary widely, from lethal to persistent infections (Bonning and Miller, 2010; van Oers, 2010). Persistent infection refers to the co-existing of the host and its pathogen in an equilibrium where virus transmission is assured in the absence of obvious fitness costs to the host (Goic and Saleh, 2012). In general, insect picornaviruses produce apparently asymptomatic infections, although due to the high viral titer they carry it is likely that more detailed studies would report effects on their fitness. In fact, the lethal and detrimental effects of insect picornaviruses have been mainly described in beneficial insects where major attention has been taken. For instance, the infectious flacherie virus produces diarrhea in the silkworm and results in important production losses (Himeno et al., 1974; Savanurmath, 1994). Sacbrood disease in honey bee larvae is a fatal disease caused by the Sacbrood virus (Bailey et al., 1964). In addition, these viruses have also been reported to produce indirect effects on the host fitness. The Kakugo virus, a subtype of the Deformed wing virus (DWV) has been found in the brain of aggressive worker honeybees.
(Fujiyuki et al., 2005) and more recently, mutualistic symbiosis between Varroa destructor, a honeybee parasite, and DWV has been reported and associated with the effect of the virus on the host immunity (Di Prisco et al., 2016).

In this work, sequence information available at the NCBI was used to identify and characterize three novel picornaviruses infecting C. capitata. Secondly, we measured the incidence of these viruses in collected field individuals and in individuals from the colonies used for mass rearing production in SIT programs. Thirdly, the possible fitness costs correlation with the viral titers were investigated. Finally, the relevance and implications of these viruses in medfly management is discussed.

2. Methods

2.1. Virus identification and genome sequence analysis

Viral sequences were identified by tBlastx against C. capitata sequences in NCBI (TSA database) using as a query the functional domains of three different RNA viruses recently described in the lepidopteran S. exigua (Jakubowska et al., 2014). Those sequences producing a significant hit were manually analyzed and filtered to remove redundant sequences. The different sequences were inspected manually in order to remove those sequences with similarity to functional domains (zinc finger domain, helicase domain, and RNA binding domain) present in viral but also in non-viral proteins, and possibly not representing viral sequences. The genomic structure, gene contents, and localization of the conserved motifs for the non-structural proteins (helicase, protease and RNA-dependent RNA polymerase), were obtained by comparison with the closest viral species.

2.2. Phylogenetic analysis

The phylogenetic analysis included most of the members of the Iflaviridae family, a few representative viruses from the Dicistroviridae family, and the unclassified viruses described
for *Drosophila* spp. and *S. exigua* (see Table S1). Multiple sequence alignment of the predicted conserved domains of the RNA-dependent RNA polymerase (RdRp) (domains I to VIII) and of the helicase (A to C motifs) from the different viruses was performed using COBALT software (Papadopoulos and Agarwala, 2007) with the RPS-BLAST and the MUSCLE alignment tool (Edgar, 2004) for the best adjustment in the alignment. The best fitting models of molecular evolution were calculated by the ProtTest2.4 software (Abascal et al., 2005).

Bayesian phylogenetic analyses were performed with the predicted RdRp and helicase domain separately using BEAST v1.8.2 (Drummond et al., 2012) with the LG+G4+I model (Le and Gascuel, 2008), and with a chain length of 10 million generations and sampling every 1000 trees, in order to establish convergence for all parameters. The BEAST outputs were analyzed using TRACER v1.5 (tree.bio.ed.ac.uk/software/tracer). The tree samples were summarized into the maximum clade credibility (MCC) phylogeny using TREEANNOTATOR v1.8.2 (beast.bio.ed.ac.uk/TreeAnnotator), discarding the first 25% of the sampled trees as burn-in. The Bayesian analyses were also performed with Dayhoff (Dayhoff, 1976), JTT (Jones et al., 1992), and BLOSUM62 (Henikoff and Henikoff, 1992) matrix resulting in the same tree structure, confirming the final structure.

### 2.3. Insects

Viral detection and abundance were determined in wild and reared adults of medfly. Wild *C. capitata* adults were collected from commercial citrus orchards located in the province of Valencia (Valencia, Spain), whereas reared *C. capitata* insects were kindly provided from a SIT mass-rearing facility (Caudete de las Fuentes, Valencia, Spain). Reared adults from three different temperature-sensitive lethal (tsl) genetic sexing strains (GSS) and one wild-type strain were used:
• **V8**) Vienna-8 mix 2002 strain [also named GS1/D53 or T(Y; 5–30C) (Franz, 2002)] (Franz, 2005), under production at the mass-rearing facility from 2007 to 2014. The strain is being reared at mass-rearing facilities for experimental goals.

• **V8G**), Vienna-8 strain lacking the D53 inversion (2013) and currently under production at mass-rearing facilities in Guatemala.

• **V8A**) a naturalized V8-strain obtained by crossing the Vienna-8 mix 2002 strain with wild individuals collected in citrus orchards located in the province of Valencia (Spain). This strain is currently under production at mass-rearing facilities in Valencia (Personal communication; Jaime García De Oteyza, Tragsa, Spain).

• **wt**) a wild strain obtained from a medfly colony maintained at the facilities of the Instituto Valenciano de Investigaciones Agrarias (IVIA). This colony was established in June 2015 with 25,000 wild medfly pupae obtained from infested fruits collected from the Navelina and Satsuma Owari abandoned orchards.

In addition, the *C. capitata* laboratory strain used in Callas et al., (2014) is related to the Vienna-8 strain and is abbreviated as V8C in this manuscript.

For the fitness-costs studies only sterilized males from the V8A and V8 strains were tested. The irradiation treatment to sterilize males was applied two days before emergence under hypoxia at 105 ± 10 Gy dose in an electron accelerator plant (IONISIOS S.A.) located in Tarancón (Cuenca, Spain). Before the beginning of the test, the two-day-old males were exposed for 3 h to an aromatherapy treatment with Ginger Root Oil (GRO) [*Zingiber officinale* Roscoe (Zingiberaceae)] (Lluch Essence S.L., El Prat de Llobregat, Barcelona, Spain) by impregnating a piece of filter paper with 2 μl of GRO (dose of 310 μl/m³) for 40 min at 3ºC to simulate the usual procedures before releasing (Juan-Blasco et al., 2014). For the mating performance tests, adults of the *wt* strain were used. All of the insects used to
determine viral detection and abundance were individually collected and conserved in 70% ethanol until processing for RNA extraction.

For mating performance tests, wild adults (hereafter wt) were obtained from a medfly colony maintained at the facilities of the Instituto Valenciano de Investigaciones Agrarias (IVIA) in a controlled environment cabinet (25 ± 1°C, 75 ± 5% RH) under illumination by fluorescent tubes (Sylvania F-18W/Grolux; 16 h day−1; 2500 lx). All of the insects derived from the mating performance tests and used for viral titering were individually collected and conserved in ethanol 70% until processing for RNA extraction.

2.4. Viral detection and quantification

Viral abundance in samples used for the transcriptomic analysis of *C. capitata* and deposited in the NCBI databases was determined by mapping of the SRA reads (Acc. number: SRS449770-72, SRS449774-75, SRS449779-80, SRS449782-92) against the genomic sequence of each of the three viruses using the Geneious for RNASeq mapper (Biomatters Ltd.). Mapping was individually performed for each of the SRA files, representing libraries derived from different tissues, conditions, or biological replicates (Calla et al., 2014). Viral abundance was obtained as RPKM (Reads Per Kilobase of transcript per Million mapped reads) values.

Presence and abundance of the RNA viruses in laboratory and field-captured insects was determined by the detection of viral RNA genomes using reverse transcription quantitative real-time polymerase-chain reaction (RT-qPCR). For this purpose, total RNA was isolated using RNAzol reagent (MRC Inc., Cincinnati, OH) according to the manufacturer’s protocol. 0.5 µg of each RNA was reverse transcribed to cDNA using random hexamers and oligo(dT) primers and following the instructions provided in the PrimeScript RT Reagent Kit (Perfect Real Time from Takara Bio Inc., Otsu Shiga, Japan). RT-qPCR was carried out in a StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA). All reactions
were performed using 5x HOT FIREpol EvaGreen qPCR Mix Plus (ROX) from Solis BioDyne (Tartu, Estonia) in a total reaction volume of 25 µl. Forward and reverse primers were designed using the Primer Express software (Applied Biosystems, Foster City, CA), and their efficiencies were calculated to be about 100% for CcaIV2 and CcaNV. In addition, the ribosomal gene \( L23a \) (Genbank acc: XM004518966) was used as an endogenous control of the RNA concentration in each RT-qPCR (Arouri et al., 2015). A list of primers used in this study is provided in Table S2. As a preliminary outcome of the RT-qPCR, the identity of the amplified fragments was confirmed by Sanger sequencing. For relative viral load quantification, Ct values from the RT-qPCR were compared to the standard curves obtained for a known number of copies of viral genome fragments cloned in the pGEMTeasy vector. Fragments of the CcaIV2 and CcaNV genomes were cloned into the pGEMTeasy vector and the standard curve prepared from the serial dilutions of known copies of the vector DNA. Viral load was normalized according to the amount of total RNA used for the cDNA synthesis. The limit of detection for each virus was established based on the average Ct values obtained for the negative controls (total RNA without cDNA synthesis). The Ct values equal or greater to the average Cts for the controls were considered free of CcaNV or CcaIV2.

Semi-quantitative RT-PCR was used to detect the negative RNA strands of CcaNV and CcaIV2 in adults. Tagged primer was used for the specific synthesis of cDNA due to the occurrence of self-priming, which is often observed for RNA viruses. RNA was extracted as described above. For this, 0.5 µg of RNA were used for cDNA synthesis using a tagged specific primer. cDNA synthesis was performed using the PrimeScript RT reagent kit from Takara Bio Inc (Otsu Shiga, Japan) following the manufacturer’s protocol, at 42ºC for 30 min. 1 µl was used for subsequent PCR reactions using the tag region as a forward primer, and the reverse primers utilized for the RT-qPCR detection of the viruses were used as
reverse primers. Negative strand detection primers are also listed in Table S2. PCR was performed using the following conditions: 95°C for 5 min, annealing at 50°C, and elongation at 72°C for 40 cycles with DreamTaq DNA polymerase (Thermo Fischer Scientific, Waltham, MA USA). The resulting PCR products were 417 bp (CcaNV) and 391 bp (CcaIV2).

2.5. Fitness studies

The possible association of the CcaNV and CcaIV2 abundance with fly fitness were studied by measuring three biological parameters following the standards of FAO/IAEA/USDA (FAO/IAEA/USDA, 2014) in V8A sterile males as easily measurable and representative of C. capitata’s fitness status:

1) Survival under stress. One hundred flies were collected after chilling (3°C for 40 min) and placed in Petri dishes (150 mm in diameter) with an opening of approximately 100 mm in the center of the lid which was fitted with a fine (16 mesh) screen for ventilation. The dishes were maintained at 25±1°C in darkness until the end of the experiment. Dead flies were removed hourly through a hole with a small stopper in the side of the Petri dish, taking care not to allow live flies to escape. The time in which each fly was removed was registered. For the viral abundance analysis, 20 dead flies were selected, the 10 with the shortest lifespan (4-7 h) and the 10 with the longest lifespan (96-125 h). Given the differences found in CcaNV abundance between groups, a second and independent batch of insects was additionally selected and analyzed as above. The abundance values reported are the combination of the two independent experiments performed in a side-by-side manner.

2) Mating performance. Adult wt females, wt males, or V8A sterile males were separated by sex and strain into perspex cages (20 x 20 x 20 cm) and kept apart in different rooms to prevent any pheromone and essence effects. Adults were maintained at 25 ±
4°C, 75 ± 5 % RH, and 14:10h (L:D) photoperiod in an environmental chamber.

Female wt adults were fed with a mixture of sugar and hydrolyzed yeast (Biokar Diagnostics Co., Pantin, France) (4:1; w:w), whereas males (wt or sterile) were fed only with sugar. Water and diet was provided ad libitum. To ensure that all of the individuals were reproductively ready to mate when the mating test was initiated, females were 10 d old, wt males 7 d old, and sterile Vienna-8 males 3 d old (Cáceres et al., 2007; Shelly, 2012). The mating arena consisted of perspex cages (30 x 40 x 20 cm) with ventilation openings. Males (30 wt and 30 V8A) were placed in the mating arena first and left to settle for 15 min, then 30 wt females were introduced, comprising the final ratio (1:1:1) tested. In each arena, the mating pairs were checked continuously for 3 h and collected in vials (50 ml volume). After copula completion (Taylor et al., 2000), the male type in each pair was determined by the presence or absence of fluorescent dye (Klassen, 2005). Twenty flies were selected from 68 sterile collected pairs and classified in early copulation (the ten first males to be collected, 6-40 minutes since the beginning of the selection), or late copulation (the ten final males to be collected, after 86-143 minutes). Ten males were selected from the V8A males remaining in the cage and classified as no copulation.

3) Flight performance. Two samples of approximately 100 chilled V8A males per sample were transferred to individual Petri dishes on which a black tube (9 cm outside diameter, 3 mm thick walls, 10 cm high) was placed on top of each Petri dish inside a ventilated Plexiglass arena (such as a 30 x 40 x 30 cm ventilated cage). Before each use, the inside of the tube was lightly coated with unscented talcum powder to prevent the flies walking out. The tubes were tapped on a firm surface to remove excess talc. The flies which had escaped from the tubes were collected from the cage, and the time since the beginning of the test was noted. For viral abundance determination 20 flies
were selected and classified into 2 groups: early flight (started flying in the first 5 minutes) or late flight (started flying between 150 and 180 minutes).

In addition, a sample of the V8 strain was also analyzed for survival under stress conditions. The flies were irradiated, GRO treated, and chilled as described for V8A males. For each of the individuals, RNA was extracted from the whole fly and the viral abundance was calculated as above. Statistical comparison among the groups was performed using the non-parametric Mann Whitney test (GraphPad Prism version 5).
3. Results

3.1. Data mining and virus identification

RdRp domains from the three *S. exigua* picornaviruses described so far were used in a tBlastn analysis against the *C. capitata* TSA databases (NCBI) to retrieve sequences potentially representing new RNA viruses. Comparison of the different obtained hits rendered 3 new sequences of more than 10 Kb long with similar genomic structure to other members of the order Picornavirales, representing three potential viruses. Detailed analysis of the genomic sequences revealed the absence of mutations or indels generating non-sense codons, and suggested that all three sequences were representative of novel viruses infecting *C. capitata*. Based on their sequence similarity to other invertebrate-infecting viruses, they were named as *C. capitata* Iflavirus 1 (CcaIV1), *C. capitata* Iflavirus 2 (CcaIV2), and *C. capitata* nora virus (CcaNV) (Fig. 1). The monopartite positive-sense single-stranded RNA (ss(+)) genomes of CcaIV1, CcaIV2, and CcaNV have lengths of 10373, 10205, and 11866 nt excluding poly(A) tails, respectively (GenBank acc: GAMC01001920.1, GAMC01020602.1, and GAMC01015827.1 for CcaIV1, CcaIV2, and CcaNV, respectively).

3.2. Genome structure and phylogenetic classification

The monocistronic genomes of CcaIV1 and CcaIV2 comprise one large uninterrupted ORF, that translates into 3068 aa and 3031 aa polyproteins, respectively. Sequence comparison between these two viruses revealed an identity of 56% at the genomic level and a similarity at the aa level of 65%. Structural and functional proteins and their cleavage sites were identified based on the homology with other iflaviruses. The N-terminal site is composed of the structural proteins VP1-VP4, headed by a leader peptide (L), while the C terminal part contains three functional proteins: helicase, protease, and RdRp (Fig. 1).

The CcaNV genomic structure diverges from the CcaIV1 and CcaIV2 genomes. It is composed of a policistronic ss(+) RNA genome containing four ORFs. In contrast to the
iflavirus genomes, the N-terminal portion of the CcaNV genome is composed of non-structural proteins, while the C-terminus contains the structural proteins. As observed for the *Drosophila* Nora virus and the *Spodoptera exigua* Nora virus helicase, protease, and RdRp constitute the ORF2 (Fig. 1), and the ORF4 and ORF5 encode structural proteins, with homology to the *Drosophila* Nora virus VP4 capsid protein. ORF1 and ORF3, positioned along the genome similarly correspond to ORFs of the *Drosophila* Nora virus.

Two functional domains encoded in all *Picornavirales*, helicase and RdRp, show sufficient sequence conservation to investigate the phylogenetic relationship between the identified viruses. According to the phylogenetic reconstruction, CcaIV1 and CcaIV2 clearly belong to the family *Iflaviridae* in the order *Picornavirales* (Fig. 2). Despite the relatively low similarity at the aa level (65%) the two Iflaviruses clustered together in the same branch. In contrast to CcaIV1 and CcaIV2, CcaNV belongs to a distantly related branch, independent from the *Iflaviridae* and *Dicistroviridae* families. CcaNV branched together with the Nora viruses from *Drosophila* spp and the Nora virus described in the Lepidoptera *S. exigua*.

3.3. Viral abundance in laboratory populations.

The three viral genomes were initially discovered in the transcriptome of *C. capitata* derived from different developmental stages, populations, and treatments (Calla et al., 2014). The transcriptome was assembled from RNAseq experiments (Illumina Hiseq 2000 sequencing) in adults and pupae from the laboratory colony, and also from field captures in Hawaii (Calla et al., 2014). The laboratory insects used in Calla et al., (2014) were derived from the V8 strain and are referred to as V8C in this work. Those experiments also included samples of γ-irradiated insects (adult and pupae), as used for the sterilization process in the SIT operational programs. The abundance of each virus in each of the samples was determined by mapping of the RNAseq reads (three replicates per sample) to the three viral genomes (Fig. 3). According to this analysis, CcaIV2 was the most abundant and ubiquitous virus, being present in all 12
samples from the V8C strain that were used in the RNAseq experiment (Fig. 3A). CcaNV was less abundant than CcaIV2, but also present in most of the V8C samples (7 out of 12). In contrast, CcaIV1 was absent in most of the samples and was only found, at very low abundance, in the three replicates of the irradiated adults of the V8C strain.

In addition, the presence of these viruses was also tested using the expression data from other C. capitata samples in the NCBI database. Although CcaIV1 and CcaNV were not detected in any of the samples available in NCBI, CcaIV2 was found in all of the available samples. This included adults and embryos from the Ispra strain (Genbank acc. SRX272876-78) and embryos of the Benakeion strain (Genbank acc. SRX591266). The viral abundance was relatively high and similar for all of the samples analyzed with values which ranked from 60000-95000 RPKM. No sequences with homology to these viruses were found in the genome of C. capitata (Papanicolaou et al., 2016), supporting the viral nature of these sequences.

The abundance of those viruses was also determined in insects from another colony which is also employed for mass production in SIT programs, the Vienna 8 strain (V8). Individual adults from that colony reared at different locations and with different treatments were tittered by qRT-PCR for the presence of the three viruses. More than 200 individuals were negative for the presence of CcaIV1 using two different sets of primers. We did not have access to insects carrying CcaIV1 as a positive control, but the absence of CcaIV1 reads in most of the samples in the NCBI database, together with the lack of amplification by qRT-PCR using two independent sets of primers, suggests that those insects are CcaIV1-free. In contrast, CcaIV2 (Fig 4A) and CcaNV (Fig 4C) were present at high titters in all of the V8-derived samples. CcaIV2 was found at a high titer in 100% of the individuals from the V8G and V8A strains. For the V8 insects, about 75% of the insects were carrying CcaIV2. In the case of CcaNV, not all of the V8-derived insects gave a positive result under our
experimental settings. The virus was detected at a greater frequency in V8A insects (90% of the tested individuals) than in the V8 and V8G samples (about 30%). Although the viral titer was greater in the V8A and V8G samples in females compared to males, in the case of the V8 insects the situation was the opposite with a higher titer in males than in females.

3.4. Viral abundance in field populations.

RNAseq samples in NCBI also included *C. capitata* samples derived from captures obtained in the field (Hawaii). Although at different levels, reads mapping to the three viruses were found for all of the adult and pupal samples. Again, CcaIV2 was the most abundant virus, followed by CcaNV and CcaIV1 (Fig 3B). Interestingly, the presence of CcaIV1 was relatively low for all of the samples tested except for a single adult sample which showed very high levels, suggesting an obvious infection (symptomatic or not) of some of the insects of that sample with CcaIV1.

We additionally tested for the presence of the three viruses in insects captured from orange orchards in three different regions in the Comunitat Valenciana (Spain) (Fig 4B and D). Again, all of the tested samples were negative for CcaIV1. In contrast, CcaIV2 (Fig 4B) and CcaNV (Fig 4D) were found in all three of the locations, being present in both males and females. The frequency of insects carrying the viruses was similar for the three locations with values of about 75% for CcaIV2 and about 50% for CcaNV.

3.5. Viral incidence in laboratory and field insects

CcaIV2 and CcaNV were simultaneously found in all of the V8-derived insects. The V8C samples were obtained from a pool of insects, and the simultaneous presence of both viruses in single insects cannot be determined. However, in the case of the V8-derived and field captured insects that were analyzed individually, the incidence of simultaneous infection was assessable. Only females were considered in order to prevent the influence of the males derived from the V8A strain captured in the field. About 60% and 40% of the females from
the laboratory and field insects, respectively, were found to be simultaneously infected with both viruses (Fig 5A). These differences could be mainly attributed to the lower incidence of CcaIV2-infected insects in the field samples. In connection with that, no laboratory insects exclusively infected with CcaNV were found, while about 20% of the field insects were only infected with that virus. In general, the viral frequency patterns between the laboratory and field insects were similar, and only about 5% of the tested insects were considered virus-free in the laboratory and field samples. Interestingly, the viral abundance was for both viruses about one order of magnitude greater in the infected females derived from the laboratory than in the infected females from the field (Fig 5B).

3.6. Viral replication assessment

A positive-strand RNA virus produces an intermediate negative-strand RNA when it replicates. Thus, the detection of negative-strand viral RNA is indicative of viral replication and supports its viral nature. Using specific PCR for the negative strand, the replication of CcaIV2 and CcaNV were observed in the larvae, pupae, and adults from the V8A strain (Fig 6).

3.7. Relationships between fly fitness and viral abundance

Due to the difficulties in obtaining virus free insects for the evaluation of the viral effects on insect fitness, we decided to analyze the correlation between the viral abundance in irradiated male adults. Initially, about 20-30 males from the V8A colony were analyzed for each parameter utilizing replicates in case significant differences for a given parameter were found. No statistically significant differences in the abundance of CcaIV2 were observed among the different groups with respect to mating performance, flying test, or lifespan (Fig. 7A and C, Fig 8A). For CcaNV, no differences were observed between groups for the mating performance and the flying test (Fig 6B and D). However, a significant difference in the CcaNV titer was observed between insects from the short life-span and long lifespan groups
The CcaNV titer was on average about 100-fold more abundant in the group with a shorter lifespan. Irradiated males from the V8 colony were also grouped according to their lifespan. In this case, differences between the groups were smaller than in the V8 insects (40 hours difference in the mean survival time). As expected, no differences in CcaIV2 abundance were observed for both groups (Fig 8B). For these groups of insects, CcaNV abundance was greater than previously observed for the V8A insects (about 100-fold), and no statistically significant differences were observed in the CcaNV titer between both groups (Fig 8D).

4. Discussion

RNA viruses comprise a large group of infectious viruses inducing diseases in plants, vertebrates, and invertebrates. The advent of genomics has accelerated the discovery of novel viruses in all types of hosts, including insects (Lecuit and Eloit, 2013; Liu et al., 2011). Here, by means of DNA sequence mining we have identified three small RNA viruses infecting an agronomically important pest *C. capitata*. Early in the 1980’s, Plus and colleagues (Plus et al., 1981), reported the presence of RNA viruses in the ISPRA strain (established in Pavia in 1979), one of them morphologically resembling picornavirus particles. It is likely that this viral preparation was composed of one or more of the viruses described here. The three viruses reported here belong to the order *Picornavirales* (Le Gall et al., 2008), and two of them clearly fall within the family *Iflaviridae*. Until this work, the *Iflaviridae* family was composed by viruses infecting insects from the orders Lepidoptera, Hemiptera, and Hymenoptera (van Oers, 2010). CcaIV1 and CcaIV2 represent the first two iflaviruses infecting dipterans. An expressed sequence tag (EST) with homology to these viruses has also been identified in another dipteran, the tsetse fly (*Glossina morsitans* virus-like sequence), suggesting that this viral family will be further expanded with new members infecting other species of dipterans. CcaNV differs from the two iflaviruses and shares
homology and genomic structure with the Nora virus from *Drosophila* spp. So far these viruses are taxonomically unassigned, but according to the phylogenetic clustering they will most likely define a new family of invertebrate viruses within the order *Picornavirales*.

The presence of CcaIV2 as well as CcaNV has been detected in most of the insects and samples analysed, representing different laboratory colonies as well as field captures. In contrast, CcaIV1 was only found in a few samples. The relatively low abundance of CcaIV1 found in samples from irradiated males from the V8C strain could suggest that male sterilization by irradiation could be contributing to the infection and/or replication of this virus. However, the virus was absent in irradiated pupae from the same colony, revealing additional or alternative factors involved in CcaIV1 infectivity and abundance. Radiation has a negative effect on an insect’s immunity, suppressing its defences against pathogens and parasitoids (Hendrichs et al., 2009), and possibly promoting viral replication. In agreement with that, although not in a side-by-side comparison, irradiated males from the V8A colony showed on average about a 100-fold increase in the viral abundance of CcaIV2 and CcaNV when compared with its non-irradiated counterparts (Fig 4A, C). Moreover, the irradiated V8 insects utilized in the lifespan studies exhibited much higher viral titers and incidence (Fig 7B, D) than previous samples of non-irradiated insects from the same colony (Fig 4A, C). Nevertheless, in the absence of side-by-side studies, and due to the different parameters that can influence the viral titer and the insect’s immune defense, the differences observed here could be the result of natural fluctuation along the generations.

Although CcaIV2 and CcaNV were found in laboratory colonies as well as in field insects, the abundance and incidence of both viruses was in general lower in field captures than in laboratory individuals. These differences could be indicative of the detrimental effects of these viruses on the insect fitness. In the absence of negative effects associated with the viral infections, it would be expected to find the viruses in most of the individuals, however on
average only about 70% and 60% of the field captures were considered positive for CcaIV2 and CcaNV, respectively. Similarly, although most of the laboratory colonies from the lepidopteran *S. exigua* are covertly infected with two species of iflaviruses (Jakubowska et al., 2014), only about 10% of the field captures were positive for any of the two iflaviruses (Virto et al., 2014).

Opposite to the relatively high persistence of the CcaIV2 and CcaNV viruses in the field, CcaIV1 was not detected in any of the individuals captured in the field in Spain. Unfortunately, no positive control was available for CcaIV1 infection and the lack of CcaIV1-infected insects could be attributed to methodological reasons. To minimize such risk, two independent sets of primers were designed to examine for the presence of this virus. Nevertheless, the failure to detect CcaIV1 could be attributed to differences in the geographical distribution of the virus, with CcaIV1 being absent in Spain. Additional screening of insects captured from different geographical locations would contribute to determining the worldwide incidence of these viruses. As mentioned above for the other viruses, the absence of CcaIV1 in the field could also reflect stronger pathological effects associated with its infection. Only 1 out of the 6 samples derived from field captures in Hawaii exhibited high levels of CcaIV1. Given the common origin of the insects in the 6 samples (Calla et al., 2014), this could be the result of the presence of strongly infected individual/s in this sample.

Two of the three novel RNA viruses were found in high titters in insects from the colonies currently in use in different SIT programs around the world. Males from these colonies are mass-produced and continuously released into the field to compete with wild males and prevent female reproduction. As a consequence, any phenomena affecting the fitness of those males can have a strong impact on the efficacy of the SIT programs (FAO/IAEA/USDA, 2014). Preliminary results have shown that those insects with lower CcaNV titers exhibit...
increased survival times, suggesting that this virus has a negative impact on the adult lifespan. Previous studies on *D. melanogaster* have also shown a decrease in the lifespan of the insects infected with DmNV when compared to the uninfected insects (Habayeb et al., 2009). Correlations between CcaNV titers and fly lifespan were found in two independent batches of the V8A colony. In addition, a new batch of the irradiated insects from the V8 colony was analysed for viral titer and lifespan, finding that those insects had in general a short lifespan and high CcaNV titer. It is tempting to speculate that the high titer found could be responsible for the shorter lifespan of those insects, although alternative hypotheses about why viral titers were higher in flies with short lifespans cannot be excluded.

Given that those viruses were able to replicate at very high titers in most of the flies and to a lower incidence and titer in the field, it is expected that more detailed studies may reveal additional sub-lethal effects of those viruses on larval and/or adult fitness. So far, the main limitation to conducting experiments to quantify in detail the viral effect on medfly fitness is the absence of virus free insects. According to the data presented here, most of the medfly stocks are persistently infected with one or even two viruses. A systematic screening on medfly colonies would contribute to the identification of those virus-free insects, and their fitness parameters when infected with individual or mixed viruses could be compared. Alternatively, certain strains of the obligate endosymbiont *Wolbachia* were shown to provide insect protection against infections with DNA and RNA viruses (including in *D. melanogaster*) (Hedges et al., 2008; Moreira et al., 2009; Teixeira et al., 2008). It would be interesting to test the presence and abundance of these viruses in *Wolbachia* infected insects (Zabalou et al., 2009; Zabalou et al., 2004). In addition, several studies evaluating the effects of microbiota composition on the fitness of SIT males have been conducted (Ami et al., 2009; Augustinos et al., 2015). It is likely, as occurred previously for other insect-virus
combinations (Jakubowska et al., 2016; Robinson and Pfeiffer, 2014), that gut microbiome composition and load would have an effect on viral replication.

Here we have described three novel viruses infecting *C. capitata*, two of them being present in most of the laboratory colonies and, although in lower frequency and titer, in field captures of this pest. Persistent infections with other insect picornaviruses have been shown to influence insect physiology and behaviour (van Oers, 2010), and the *C. capitata* viruses reported in this study may similarly affect fitness of adult flies. Independent of the lack of conspicuous effects of these viruses on laboratory-reared insects, they may play an important role in the pest biology and dynamics in the field. These types of viruses may influence, directly or indirectly, the insect’s immune system, as well as the susceptibility of this pest to other microbial pathogens (Jakubowska et al., 2016) or parasites (Di Prisco et al., 2016), as occurs for other insect-virus combinations. Additional studies evaluating the consequences of covert infections with RNA viruses in medfly, a model species for SIT applications, would shed light on the applied and ecological roles of these viruses, and contribute to increasing the efficacy of the control of this pest by better SIT approaches, or even by the use of these viruses as pest control agents.

5. Acknowledgments:

We want to thank Jaime Garcia De Oteyza and Ignacio Pla Mora from TRAGSA for providing us with the laboratory and field insects. This study was partially supported by the Spanish Ministry for Science and Technology (AGL2014-57752-C2) and by European FEDER funds. AL-G was recipient of a PhD grant from the Spanish Ministry of Science and Innovation (grant ref. BES-2015-071369). MC was recipient of a specialization grant from the European Social Fund (FSE-CV-2007-2013).
Figure legends

Fig. 1. Schematic representation of the genome organization for the three C. capitata picornaviruses. Structural proteins: VP1, VP2, VP3, and VP4. UP1 and UP3 are proteins in CcaNV of unknown function. Non-structural (functional) proteins: Hel-helicase, Pro-protease, and RdRp-RNA dependent RNA polymerase; * predicted protease cleavage sites of the capsid proteins, VPg a small < 5 kDa virus protein 5’ bound; 5’ and 3’UTR (untranslated regions); AAAn-poly(A) tail. All genomes are drawn proportionally, and the nucleotide kilobase scale is included at the bottom. Homologous proteins are marked with the same color.

Fig 2. Phylogenetic trees based on Bayesian analysis of the conserved amino acid sequences containing domains I to VIII of the RdRp and A to C of the Helicase from different members of the order Picornavirales. Amino acid sequences were aligned using the COBALT and MUSCLE softwares. Phylogenetic relationships were reconstructed using BEAST software with a LG+GI model. Posterior probability is indicated in branches (that are at least 0.60). The scale bar indicates an evolutionary distance of 0.2 amino acid substitutions per position in the sequence. The order of the host species is shown in brackets (Hy: Hymenoptera, He: Heteroptera, D: Diptera, L: Lepidoptera, Or: Orthoptera). The GenBank accession numbers of the sequences employed in the analysis are listed in supplementary Table (S1).
**Fig. 3.** Viral abundance obtained by mapping of the RNAseq against the three viruses in the different samples used for the transcriptome of *C. capitata* (Calla et al., 2014). Abundance is measured as RPKM. Each spot represents the values for each of the datasets. ND is used for those samples without reads mapping into the viral sequences.

**Fig. 4.** Relative abundance of viral RNA in *C. capitata* adults from different colonies used in SIT programs and field captures obtained by RT-qPCR. CcaIV2 abundance in insects from the V8, V8G, and V8A colonies (A) and from insects captured in different geographic locations (B). CcaNV abundance in insects from the V8, V8G, and V8A colonies (C) and from insects captured in different geographic locations (D). Each spot represents the value for an individual. ND (non-detected) is used for those samples under the detection threshold. The percentage of positive individuals is reported in each sample.

**Fig. 5.** Incidence of viral infection in laboratory and field-captured insects. A) Frequency of females infected with each of the viruses (individually and simultaneously) in laboratory and field samples. B) Average viral load for CcaIV2 and CCaNV in laboratory and field-captured females.

**Fig. 6.** Detection of the viral negative strand intermediate. Specific amplification of the negative strand of CcaIV2 (upper panel) and CcaNV (lower panel) in larvae, pupae, and adults from the V8A colony.

**Fig. 7.** Relative abundance of viral RNA in *C. capitata* males from the fitness studies obtained by RT-qPCR. CcaIV2 abundance in males from the V8A strain grouped according their flying response (A) or mating behavior (B). CcaNV abundance in males from the V8A strain grouped according to their flying response (C) or mating behavior (D). Each spot represents the value for an individual. ND (non-detected) is used for those samples under the detection threshold.

**Fig. 8.** Relative abundance of viral RNA in *C. capitata* adult males from the fitness studies obtained by RT-qPCR. CcaIV2 abundance in males from the V8A (A) and V8 (B) colonies grouped according to their lifespan. CcaNV abundance in males from the V8A (C) and V8 (D) colonies grouped according to their lifespan. Each spot represents the value for an
individual. The abundance values reported are the combination of two independent and side-by-side experiments. Statistically significant differences were only found for CcaNV in two replicates for the V8A comparisons (P-values of 0.003 and 0.05 for each of the replicates). ND (non-detected) is used for those samples under the detection threshold. MTS (mean time of survival) is reported for each group.
References


Virto, C., et al., 2014. Natural populations of Spodoptera exigua are infected by multiple viruses that are transmitted to their offspring. Journal of Invertebrate Pathology. 122, 22-27.

Figure 6
Click here to download high resolution image

![Image of gel electrophoresis results for CcaLV2 and CcaNV](image-url)
Figure 8
Click here to download high resolution image
Table S1
Click here to download Supplementary Material: TableS1.docx
Table S2
Click here to download Supplementary Material: TableS2_R1.docx